

*re-run*

*#13*

## **RAW SEQUENCE LISTING**

**The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) no errors detected.**

Application Serial Number: 09/554,533A  
Source: PT  
Date Processed by STIC: 5/17/05

# ***ENTERED***



PCT

## RAW SEQUENCE LISTING

DATE: 05/11/2005

PATENT APPLICATION: US/09/554,533A

TIME: 18:13:58

Input Set : N:\AMC\US09554533A.raw

Output Set: N:\CRF4\05112005\I554533A.raw

C--> 1 <110> APPLICANT: BEELEY, NIGEL R. A.  
2 PRICKETT, KATHRYN S.  
3 <120> TITLE OF INVENTION: NOVEL EXENDIN AGONIST COMPOUNDS  
4 <130> FILE REFERENCE: 18528.016 (238/086 US)  
5 <140> CURRENT APPLICATION NUMBER: US/09/554,533A  
6 <141> CURRENT FILING DATE: 2000-05-11  
7 <150> PRIOR APPLICATION NUMBER: PCT/US98/24210  
8 <151> PRIOR FILING DATE: 1998-11-13  
9 <150> PRIOR APPLICATION NUMBER: 60/065,442  
10 <151> PRIOR FILING DATE: 1997-11-14  
11 <160> NUMBER OF SEQ ID NOS: 87  
12 <170> SOFTWARE: PatentIn version 3.2  
14 <210> SEQ ID NO: 1  
15 <211> LENGTH: 39  
16 <212> TYPE: PRT  
17 <213> ORGANISM: Heloderma horridum  
18 <220> FEATURE:  
19 <223> OTHER INFORMATION: C-term amidated  
20 <400> SEQUENCE: 1  
21 His Ser Asp Gly Thr Phe Thr Ser Asp Leu Ser Lys Gln Met Glu Glu  
22 1 5 10 15  
23 Glu Ala Val Arg Leu Phe Ile Glu Trp Leu Lys Asn Gly Gly Pro Ser  
24 20 25 30  
25 Ser Gly Ala Pro Pro Pro Ser  
26 35  
28 <210> SEQ ID NO: 2  
29 <211> LENGTH: 39  
30 <212> TYPE: PRT  
31 <213> ORGANISM: Heloderma suspectum  
32 <220> FEATURE:  
33 <223> OTHER INFORMATION: C-term amidated  
34 <400> SEQUENCE: 2  
35 His Gly Glu Gly Thr Phe Thr Ser Asp Leu Ser Lys Gln Met Glu Glu  
36 1 5 10 15  
37 Glu Ala Val Arg Leu Phe Ile Glu Trp Leu Lys Asn Gly Gly Pro Ser  
38 20 25 30  
39 Ser Gly Ala Pro Pro Pro Ser  
40 35  
42 <210> SEQ ID NO: 3  
43 <211> LENGTH: 30  
44 <212> TYPE: PRT  
45 <213> ORGANISM: Homo sapien  
46 <220> FEATURE:

P.6

## RAW SEQUENCE LISTING

DATE: 05/11/2005

PATENT APPLICATION: US/09/554,533A

TIME: 18:13:58

Input Set : N:\AMC\US09554533A.raw

Output Set: N:\CRF4\05112005\I554533A.raw

47 <223> OTHER INFORMATION: C-term amidated  
48 <400> SEQUENCE: 3  
49 His Ala Glu Gly Thr Phe Thr Ser Asp Val Ser Ser Tyr Leu Glu Gly  
50 1 5 10 15  
51 Gln Ala Ala Lys Glu Phe Ile Ala Trp Leu Val Lys Gly Arg  
52 20 25 30  
54 <210> SEQ ID NO: 4  
55 <211> LENGTH: 38  
56 <212> TYPE: PRT  
57 <213> ORGANISM: Artificial Sequence  
58 <220> FEATURE:  
59 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic peptide  
60 <220> FEATURE:  
61 <221> NAME/KEY: MOD\_RES  
62 <222> LOCATION: (1)  
63 <223> OTHER INFORMATION: His, Arg or Tyr  
64 <220> FEATURE:  
65 <221> NAME/KEY: MOD\_RES  
66 <222> LOCATION: (2)  
67 <223> OTHER INFORMATION: Ser, Gly, Ala or Thr  
68 <220> FEATURE:  
69 <221> NAME/KEY: MOD\_RES  
70 <222> LOCATION: (3)  
71 <223> OTHER INFORMATION: Ala, Asp, or Glu  
72 <220> FEATURE:  
73 <221> NAME/KEY: MOD\_RES  
74 <222> LOCATION: (5)  
75 <223> OTHER INFORMATION: Ala or Thr  
76 <220> FEATURE:  
77 <221> NAME/KEY: MOD\_RES  
78 <222> LOCATION: (6)  
79 <223> OTHER INFORMATION: Ala, Phe, Tyr or naphthylalanine  
80 <220> FEATURE:  
81 <221> NAME/KEY: MOD\_RES  
82 <222> LOCATION: (7)  
83 <223> OTHER INFORMATION: Thr or Ser  
84 <220> FEATURE:  
85 <221> NAME/KEY: MOD\_RES  
86 <222> LOCATION: (8)  
87 <223> OTHER INFORMATION: Ala, Ser or Thr  
88 <220> FEATURE:  
89 <221> NAME/KEY: MOD\_RES  
90 <222> LOCATION: (9)  
91 <223> OTHER INFORMATION: Asp or Glu  
92 <220> FEATURE:  
93 <221> NAME/KEY: MOD\_RES  
94 <222> LOCATION: (10)  
95 <223> OTHER INFORMATION: Ala, Leu, Ile, Val, pentylglycine, or Met  
96 <220> FEATURE:

## RAW SEQUENCE LISTING

DATE: 05/11/2005

PATENT APPLICATION: US/09/554,533A

TIME: 18:13:58

Input Set : N:\AMC\US09554533A.raw

Output Set: N:\CRF4\05112005\I554533A.raw

97 <221> NAME/KEY: MOD\_RES  
98 <222> LOCATION: (11)  
99 <223> OTHER INFORMATION: Ala or Ser  
100 <220> FEATURE:  
101 <221> NAME/KEY: MOD\_RES  
102 <222> LOCATION: (12)  
103 <223> OTHER INFORMATION: Ala or Lys  
104 <220> FEATURE:  
105 <221> NAME/KEY: MOD\_RES  
106 <222> LOCATION: (13)  
107 <223> OTHER INFORMATION: Ala or Gln  
108 <220> FEATURE:  
109 <221> NAME/KEY: MOD\_RES  
110 <222> LOCATION: (14)  
111 <223> OTHER INFORMATION: Ala, Leu, Ile pentylglycine, Val, or Met  
112 <220> FEATURE:  
113 <221> NAME/KEY: MOD\_RES  
114 <222> LOCATION: (15)  
115 <223> OTHER INFORMATION: Ala or Glu  
116 <220> FEATURE:  
117 <221> NAME/KEY: MOD\_RES  
118 <222> LOCATION: (16)  
119 <223> OTHER INFORMATION: Ala or Glu  
120 <220> FEATURE:  
121 <221> NAME/KEY: MOD\_RES  
122 <222> LOCATION: (17)  
123 <223> OTHER INFORMATION: Ala or Glu  
124 <220> FEATURE:  
125 <221> NAME/KEY: MOD\_RES  
126 <222> LOCATION: (19)  
127 <223> OTHER INFORMATION: Ala or Val  
128 <220> FEATURE:  
129 <221> NAME/KEY: MOD\_RES  
130 <222> LOCATION: (20)  
131 <223> OTHER INFORMATION: Ala or Arg  
132 <220> FEATURE:  
133 <221> NAME/KEY: MOD\_RES  
134 <222> LOCATION: (21)  
135 <223> OTHER INFORMATION: Ala or Leu  
136 <220> FEATURE:  
137 <221> NAME/KEY: MOD\_RES  
138 <222> LOCATION: (22)  
139 <223> OTHER INFORMATION: Phe, Tyr, or naphthylalanine  
140 <220> FEATURE:  
141 <221> NAME/KEY: MOD\_RES  
142 <222> LOCATION: (23)  
143 <223> OTHER INFORMATION: Ile, Val, Leu, pentylglycine, tert-butylglycine,  
144 or Met  
145 <220> FEATURE:

## RAW SEQUENCE LISTING

DATE: 05/11/2005

PATENT APPLICATION: US/09/554,533A

TIME: 18:13:58

Input Set : N:\AMC\US09554533A.raw

Output Set: N:\CRF4\05112005\I554533A.raw

146 <221> NAME/KEY: MOD\_RES  
147 <222> LOCATION: (24)  
148 <223> OTHER INFORMATION: Ala, Glu, or Asp  
149 <220> FEATURE:  
150 <221> NAME/KEY: MOD\_RES  
151 <222> LOCATION: (25)  
152 <223> OTHER INFORMATION: Ala, Trp, Phe, Tyr, or naphthylalanine  
153 <220> FEATURE:  
154 <221> NAME/KEY: MOD\_RES  
155 <222> LOCATION: (26)  
156 <223> OTHER INFORMATION: Ala or Leu  
157 <220> FEATURE:  
158 <221> NAME/KEY: MOD\_RES  
159 <222> LOCATION: (27)  
160 <223> OTHER INFORMATION: Ala or Lys  
161 <220> FEATURE:  
162 <221> NAME/KEY: MOD\_RES  
163 <222> LOCATION: (28)  
164 <223> OTHER INFORMATION: Ala or Asn  
165 <220> FEATURE:  
166 <221> NAME/KEY: MOD\_RES  
167 <222> LOCATION: (29)  
168 <223> OTHER INFORMATION: Gly or not present  
169 <220> FEATURE:  
170 <221> NAME/KEY: MOD\_RES  
171 <222> LOCATION: (30)  
172 <223> OTHER INFORMATION: Gly or not present  
173 <220> FEATURE:  
174 <221> NAME/KEY: MOD\_RES  
175 <222> LOCATION: (31)  
176 <223> OTHER INFORMATION: Pro, homoproline, 3Hyp, 4Hyp, thioproline,  
177 N-alkylglycine, N-alkylpentylglycine,  
178 N-alkylalanine, or not present  
179 <220> FEATURE:  
180 <221> NAME/KEY: MOD\_RES  
181 <222> LOCATION: (32)  
182 <223> OTHER INFORMATION: Ser or not present  
183 <220> FEATURE:  
184 <221> NAME/KEY: MOD\_RES  
185 <222> LOCATION: (33)  
186 <223> OTHER INFORMATION: Ser or not present  
187 <220> FEATURE:  
188 <221> NAME/KEY: MOD\_RES  
189 <222> LOCATION: (34)  
190 <223> OTHER INFORMATION: Gly or not present  
191 <220> FEATURE:  
192 <221> NAME/KEY: MOD\_RES  
193 <222> LOCATION: (35)  
194 <223> OTHER INFORMATION: Ala or not present

## RAW SEQUENCE LISTING

DATE: 05/11/2005

PATENT APPLICATION: US/09/554,533A

TIME: 18:13:58

Input Set : N:\AMC\US09554533A.raw

Output Set: N:\CRF4\05112005\I554533A.raw

195 <220> FEATURE:  
196 <221> NAME/KEY: MOD\_RES  
197 <222> LOCATION: (36)  
198 <223> OTHER INFORMATION: Pro, homoproline, 3Hyp, 4Hyp, thioproline,  
199 N-alkylglycine, N-alkylpentylglycine,  
200 N-alkylalanine, or not present  
201 <220> FEATURE:  
202 <221> NAME/KEY: MOD\_RES  
203 <222> LOCATION: (37)  
204 <223> OTHER INFORMATION: Pro, homoproline, 3Hyp, 4Hyp, thioproline,  
205 N-alkylglycine, N-alkylpentylglycine,  
206 N-alkylalanine, or not present  
207 <220> FEATURE:  
208 <221> NAME/KEY: MOD\_RES  
209 <222> LOCATION: (38)  
210 <223> OTHER INFORMATION: Pro, homoproline, 3Hyp, 4Hyp, thioproline,  
211 N-alkylglycine, N-alkylpentylglycine,  
212 N-alkylalanine, or not present  
213 <220> FEATURE:  
214 <223> OTHER INFORMATION: May be c-term amidated  
215 <400> SEQUENCE: 4  
W--> 216 Xaa Xaa Xaa Gly Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa  
217 1 5 10 15  
218 Xaa Ala Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa  
219 20 25 30  
220 Xaa Xaa Xaa Xaa Xaa Xaa  
221 35  
223 <210> SEQ ID NO: 5  
224 <211> LENGTH: 30  
225 <212> TYPE: PRT  
226 <213> ORGANISM: Artificial Sequence  
227 <220> FEATURE:  
228 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic peptide  
229 <220> FEATURE:  
230 <223> OTHER INFORMATION: May be c-term amidated  
231 <400> SEQUENCE: 5  
232 His Gly Glu Gly Thr Phe Thr Ser Asp Leu Ser Lys Gln Met Glu Glu  
233 1 5 10 15  
234 Glu Ala Val Arg Leu Phe Ile Glu Trp Leu Lys Asn Gly Gly  
235 20 25 30  
237 <210> SEQ ID NO: 6  
238 <211> LENGTH: 28  
239 <212> TYPE: PRT  
240 <213> ORGANISM: Artificial Sequence  
241 <220> FEATURE:  
242 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic peptide  
243 <220> FEATURE:  
244 <223> OTHER INFORMATION: May be c-term amidated  
245 <400> SEQUENCE: 6

RAW SEQUENCE LISTING ERROR SUMMARY      DATE: 05/11/2005  
PATENT APPLICATION: US/09/554,533A      TIME: 18:13:59

Input Set : N:\AMC\US09554533A.raw  
Output Set: N:\CRF4\05112005\I554533A.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:4; Xaa Pos. 1,2,3,5,6,7,8,9,10,11,12,13,14,15,16,17,19,20,21,22,23,24  
Seq#:4; Xaa Pos. 25,26,27,28,29,30,31,32,33,34,35,36,37,38  
Seq#:47; Xaa Pos. 31,36,37,38  
Seq#:48; Xaa Pos. 36,37,38  
Seq#:49; Xaa Pos. 31  
Seq#:50; Xaa Pos. 31,36,37  
Seq#:51; Xaa Pos. 31,36,37  
Seq#:52; Xaa Pos. 31,36  
Seq#:55; Xaa Pos. 6  
Seq#:59; Xaa Pos. 10  
Seq#:60; Xaa Pos. 22  
Seq#:61; Xaa Pos. 23  
Seq#:65; Xaa Pos. 31,36,37  
Seq#:66; Xaa Pos. 1,2,3,5,6,7,8,9,10,11,12,13,14,15,16,17,19,20,21,22,23,24  
Seq#:66; Xaa Pos. 25,26,27,28,29,30,31,32,33,34,35,36,37,38,39,40  
Seq#:67; Xaa Pos. 1,26  
Seq#:68; Xaa Pos. 1,26  
Seq#:69; Xaa Pos. 1,26  
Seq#:70; Xaa Pos. 1,26  
Seq#:71; Xaa Pos. 1,27  
Seq#:72; Xaa Pos. 1,27  
Seq#:73; Xaa Pos. 1,27  
Seq#:74; Xaa Pos. 1,27  
Seq#:75; Xaa Pos. 3  
Seq#:76; Xaa Pos. 3  
Seq#:77; Xaa Pos. 3  
Seq#:78; Xaa Pos. 3,8  
Seq#:79; Xaa Pos. 3,8,9  
Seq#:80; Xaa Pos. 3,8,9,10  
Seq#:81; Xaa Pos. 3  
Seq#:82; Xaa Pos. 3  
Seq#:83; Xaa Pos. 3  
Seq#:84; Xaa Pos. 3,8  
Seq#:85; Xaa Pos. 3,8,9  
Seq#:86; Xaa Pos. 3,8,9,10  
Seq#:87; Xaa Pos. 1,2,3,5,6,7,8,9,10,11,12,13,14,15,16,17,19,20,21,22,23,24  
Seq#:87; Xaa Pos. 25,26,27,28,29,30,31,32,33,34,35,36,37,38

## VERIFICATION SUMMARY

DATE: 05/11/2005

PATENT APPLICATION: US/09/554,533A

TIME: 18:13:59

Input Set : N:\AMC\US09554533A.raw

Output Set: N:\CRF4\05112005\I554533A.raw

L:5 M:270 C: Current Application Number differs, Wrong Format  
L:216 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:4 after pos.:0  
M:341 Repeated in SeqNo=4  
L:850 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:47 after pos.:16  
M:341 Repeated in SeqNo=47  
L:872 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:48 after pos.:32  
L:890 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:49 after pos.:16  
L:914 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:50 after pos.:16  
M:341 Repeated in SeqNo=50  
L:938 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:51 after pos.:16  
M:341 Repeated in SeqNo=51  
L:962 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:52 after pos.:16  
M:341 Repeated in SeqNo=52  
L:1010 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:55 after pos.:0  
L:1070 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:59 after pos.:0  
L:1090 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:60 after pos.:16  
L:1108 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:61 after pos.:16  
L:1173 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:65 after pos.:16  
M:341 Repeated in SeqNo=65  
L:1348 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:66 after pos.:0  
M:341 Repeated in SeqNo=66  
L:1372 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:67 after pos.:0  
M:341 Repeated in SeqNo=67  
L:1394 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:68 after pos.:0  
M:341 Repeated in SeqNo=68  
L:1416 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:69 after pos.:0  
M:341 Repeated in SeqNo=69  
L:1438 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:70 after pos.:0  
M:341 Repeated in SeqNo=70  
L:1460 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:71 after pos.:0  
M:341 Repeated in SeqNo=71  
L:1482 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:72 after pos.:0  
M:341 Repeated in SeqNo=72  
L:1504 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:73 after pos.:0  
M:341 Repeated in SeqNo=73  
L:1526 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:74 after pos.:0  
M:341 Repeated in SeqNo=74  
L:1544 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:75 after pos.:0  
L:1563 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:76 after pos.:0  
L:1582 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:77 after pos.:0  
L:1607 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:78 after pos.:0  
L:1638 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:79 after pos.:0  
L:1674 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:80 after pos.:0  
L:1689 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:81 after pos.:0  
L:1707 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:82 after pos.:0  
L:1725 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:83 after pos.:0  
L:1748 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:84 after pos.:0  
L:1776 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:85 after pos.:0



VERIFICATION SUMMARY

DATE: 05/11/2005

PATENT APPLICATION: US/09/554,533A

TIME: 18:13:59

Input Set : N:\AMC\US09554533A.raw

Output Set: N:\CRF4\05112005\I554533A.raw

L:1808 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:86 after pos.:0

L:1969 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:87 after pos.:0

M:341 Repeated in SeqNo=87